

# UCLA

## UCLA Previously Published Works

### Title

Robot life: simulation and participation in the study of evolution and social behavior.

### Permalink

<https://escholarship.org/uc/item/2295v3j2>

### Journal

History and philosophy of the life sciences, 40(1)

### ISSN

0391-9714

### Author

Kelty, Christopher M

### Publication Date

2018

### DOI

10.1007/s40656-017-0181-y

Peer reviewed

# Robot life: simulation and participation in the study of evolution and social behavior

Christopher M. Kelty<sup>1</sup>

Received: 19 December 2016 / Accepted: 20 December 2017 / Published online: 3 January 2018  
© Springer International Publishing AG, part of Springer Nature 2017

**Abstract** This paper explores the case of using robots to simulate evolution, in particular the case of Hamilton’s Law. The uses of robots raises several questions that this paper seeks to address. The first concerns the role of the robots in biological research: do they simulate something (life, evolution, sociality) or do they participate in something? The second question concerns the physicality of the robots: what difference does embodiment make to the role of the robot in these experiments. Thirdly, how do life, embodiment and social behavior relate in contemporary biology and why is it possible for robots to illuminate this relation? These questions are provoked by a strange similarity that has not been noted before: between the problem of simulation in philosophy of science, and Deleuze’s reading of Plato on the relationship of ideas, copies and simulacra.

**Keywords** Simulation · Evolution · Robots · Participation · Social behavior · Deleuze

## 1 Introduction

A 2011 paper in the journal *PLOS Biology* described an experimental test of Hamilton’s rule; the authors report success (Waibel et al. 2011). It is, they claim, “the first quantitative test of Hamilton’s rule in a system with a complex mapping between genotype and phenotype” and they “demonstrate the wide applicability of kin selection theory.”

This is surprising for many reasons. First, Hamilton’s rule is one of the more controversial features of evolutionary biology in the last half-century. Hamilton’s

---

✉ Christopher M. Kelty  
ckelty@ucla.edu

<sup>1</sup> The UCLA Institute for Society and Genetics, Box 957221, 3360 LSB, Campus Mailcode: 722105, Los Angeles, CA 90095-7221, USA

Law (often written “ $rB > C$ ”) asserts that a behavior that otherwise seems to be unselfish, perhaps even altruistic, might actually confer a fitness benefit (i.e. more success in reproduction) because it increases the chance that not only an individual’s genes, but those of his close kin, will succeed. Hamilton gave precise form to this intuition by introducing a formalism for what is now called “kin selection”: if the organism benefitting from the altruist is sufficiently genetically related to it, then it might cancel out the reproductive cost of that act to the altruist (Hamilton 1964).<sup>1</sup>

Second the authors refer to their experiment as one in “artificial evolution.” This is obvious enough given that experimenting on evolution is not easy to do with any but the most short-lived organisms like fruit flies and microbes, except that, third, this experiment did not use social organisms of any predictable kind, such as wasps or ants, or even microbes, but robots: tiny mobile robots named *ALICE*.

These robots raise a set of confusing questions that this article will address. The first concerns the role of the robots in biological research: do they *simulate* something (life, evolution, sociality) or do they *participate* in something? The second question concerns the physicality of the robots: what difference does embodiment make to the role of the robot in these experiments? In these experiments, there are some subtle distinctions between what is abstract, what is digital, what is simulated, and what is physical. Thirdly, how are life, embodiment and social behavior related in contemporary biology, and why is it possible for robots to illuminate this relation? These questions are provoked by a strange similarity that has not been noted before: between the problem of simulation in philosophy of science, and Deleuze’s reading of Plato on the relationship of ideas, copies and simulacra. Whereas robot scientists, biologists and some philosophers of science may argue that robots are an “object” on which one can do precise experiments about a “target” (life, social behavior, evolution), Deleuze might counsel instead that we do not treat robots enough like robots, nor take our curiosity about them quite seriously enough—for they themselves might be the object we are studying, or should be. If Deleuze’s solution to the problem of Plato’s metaphysics implicates questions of participation and simulation, then it ought also to have implications for what robots are, and what they can tell us about evolution, behavior, or life.

The paper tracks back and forth between cases of contemporary biology using robots, and an exploration of Deleuze’s reading of Plato and some philosophical issues related to simulation in the sciences. The first part of this paper reviews attempts to use robots in biology, and explains how the core case studied here—Floreano and Keller’s experiment—was conducted. I then introduce Deleuze’s reading of Plato, followed by a discussion of how robots are used as elements in a simulation of various aspects of biology. I then explain why there might be skepticism about the use of robots to study life, but that they fit squarely into debates about simulation in the sciences and that scientists themselves are concerned to

<sup>1</sup> The concept of kin selection has been controversial ever since—whether associated with sociobiology, evolutionary psychology or the biology of social behavior; see (Kurland 1980; Leigh 2010; Okasha 2001). In 2010 E.O. Wilson and Martin Nowak recently made themselves the object of some quite hostile and aggressive attacks in the biology community by suggesting that kin selection and Hamilton’s rule are wrong (Nowak et al. 2010).

articulate where on a “scale of situatedness” robots should be placed. In the last two sections, I explore how the physicality of robots (and the relation of physicality and sociality) is used to study differences between robots and living organisms. In the conclusion, I return to the question of whether robots simulate or participate in life, or whether they “animate” theories in a novel way. A coda returns to the speculative questions raised by Deleuze’s reading of Plato.

## 2 The robot invasion begins

Using robots to study evolution might seem at first extremely unlikely—but the practice has steadily grown more common over the last twenty years, owing in large part to the manifest enthusiasm that computer scientists and engineers have shown for evolutionary theory as both a theory and as an engineering principle. Many robot scientists, computer scientists, and artificial life researchers have captured evolutionary theory in a bottle, as it were, and found ways to explore it, exploit it and ponder it as such, without ever coming near a living entity other than, arguably, themselves (see for example Eiben and Smith 2015). Recently, however, biologists have also started to take such work a bit more seriously. Several reviews have appeared recently aimed at enlightening biologists about this work (Garnier 2011; Krause et al. 2011; Mitri et al. 2013). And indeed, one of the authors of the paper about Hamilton’s rule, Laurent Keller (University of Lausanne), is a widely-lauded biologist of social evolution (primarily in real ants) who is thus lending a modicum of legitimacy to a style of experiment that might otherwise be dismissed by his colleagues.

Since about 2005, Keller and collaborator Dario Floreano, a robot scientist at École Polytechnique Fédérale de Lausanne (EPFL), have done a number of different experiments with robots that explore aspects of evolution. In most of them, tiny robots move around on motorized wheels in an enclosed area and interact with each other. Often there is “food” or “poison” (objects with blinking LEDs for instance that can be detected by a sensor), which the robots are programmed to forage for (or to avoid). In some the robots are predator or prey. All told there is a small taxon of robots occupying the ecological niche that is Floreano’s Laboratory of Intelligent Systems (<http://lis.epfl.ch/>): “Khepera” robots, “S-bots” “Alice robots” and “E-pucks.” Properly programmed and set to work, the robots have demonstrated the evolution of communication, the evolution of altruism, the evolution of communication suppression, the relationship of signal reliability and relatedness; predator prey co-evolution; and various aspects of the morphology of a robot body (Floreano et al. 2007; Floreano and Keller 2010; Mitri et al. 2009, 2011).

In order to test Hamilton’s rule, the authors constructed an experimental arena in which robots with artificial genomes (software) competed to “forage” for food and then they either “consumed” the food, or shared it with the other robots in the arena (in groups of 8). After a set amount of time, the success of the robots at achieving this task was measured, and the outcomes determined which genomes would be transmitted to the next generation. Over many generations and across several treatments, the authors determined that Hamilton’s rule held.

The details matter here: the Alice robots had infrared sensors that allowed them to differentiate the other robots from the “food items” deposited amongst them, and they had two vision sensors that allow them to “see” the walls (three black and one white) of the enclosure. In each case a group of 8 robots were instructed to gather up the food (i.e. push it towards the white wall) and then decide whether or not to “share” it with the other 7 robots. The robots do not consume the food, nor metabolize it—they are designed only to be highly simplified phenotypic vehicles for a limited set of genes that govern their behavior.<sup>2</sup> Apparently, the life of an ALICE robot is nasty and short, but only occasionally brutish. They do not reproduce, sexually or asexually, but are manually regenerated by a combination of precise but capricious software (which combines a selection event related to foraging efficiency, a crossover event, and a source of mutation) and beneficent human intervention (the robots need to be plugged in or otherwise connected to a computer in order to download the new genome and become their own offspring). An important caveat to note here is that the robots do not ever develop—ontogenically speaking—they enter and leave each generation fully formed and unchanged; as a result there is also no inter-generational interaction as part of the robot evolutionary model in these studies.

Since these robots have neither brains nor DNA, they are instead equipped with a software-based model of the same: a “neural network” consisting of 13 “neurons” (2 input neurons, 3 output neurons and 3 hidden neurons) and the 33 connections among them. These connections were governed (in software) by 33 “genes” that gave a specific *weight* to the connection, thus determining the complex behavior of the robot based on the inputs received, and the translation of these inputs into the simple possible outputs: turn the right motor, turn the left motor, or “share” the food that had been pushed to the white wall with other robots.

This simple set-up allows the experimenters to precisely vary the measure of relatedness in Hamilton’s rule ( $r$ ), and watch as the robot’s fitness changes over a series of generations. Relatedness was created within groups of robots by cloning individuals—so a group with a relatedness ( $r$ ) of 0.5 could be achieved by taking three different individuals and cloning one of them 6 times. Fitness was defined in terms of the reward for successfully transporting food to the wall—two different kinds of rewards were given for shared and non-shared food. When a robot did not share food, it received a benefit ( $c$ ) to itself. When it shared food, a benefit ( $b$ ) was distributed equally to the seven other robots in the group. Experimentally, therefore, all three components of Hamilton’s rule ( $r$ ,  $b$  and  $c$ ) could be controlled and measured.

In each generation, the robots’ “foraging efficiency” was measured (how quickly they identified and pushed food to the wall), and the inclusive fitness of each individual was determined by how many times a robot shared the food, and how many times food was shared by another robot in the group. The robots were then “selected” based on inclusive fitness, subjected to crossover with a random partner,

<sup>2</sup> Another set of unrelated researchers has explored metabolism using robots, though seemingly only to prove that a robot thus designed can keep itself running by consuming an “anaerobic or pasteurized sludge”: (Ieropoulos 2003; Ieropoulos et al. 2010; Lowe et al. 2010). They do not share the sludge with each other as far as I can determine.

subjected to mutations, and then the next generation (8 robots reloaded with the updated software) was put back into the arena with new food items and new genomes representing their changed fitness.<sup>3</sup> After hundreds of generations, the gene frequencies of the resulting robots confirmed just what Hamilton's rule would predict: some of the robots "laid down their lives to save 2 brothers or 8 cousins."

As the experimenters put it:

Because the 33 genes were initially set to random values, the robots' behaviors were completely arbitrary in the first generation. However, the robots' performance rapidly increased over the 500 generations of selection. The level of altruism also rapidly changed over generations with the final stable level of altruism varying greatly depending on within group relatedness and c/b ratio. (Waibel et al. 2011: p. 2)

A careful reading of the article reveals that the experiment was conducted with 200 groups of 8 robots over 500 generations and 25 different treatments (5 different c/b ratios and 5 relatedness (r) ratios). That is a lot of robots. It would be 20 million robots if they were not re-usable bodies for new genomes (which they are)—but even 1600 is a lot of robots. But as the authors explain: "because of the impossibility to conduct hundreds of generations of selection with real robots, we used physics-based simulations that precisely model the dynamical and physical properties of the robots." So to be clear: there are at least two kinds of robots in these experiments: "real" (physical) robots and simulated robots.<sup>4</sup> But both kinds do "exactly" the same thing, which raises the interesting question: why use physical robots at all? What exactly is the difference between a physical robot and a simulated robot? Do robots *simulate* nature or *participate* in it?

### 3 I'm not a robot, but I play one on TV...

*In Plato's world variation is accidental, while essences record a higher reality; in Darwin's reversal, we value variation as a defining (and concrete earthly) reality, while averages (our closest operational approach to essences) become mental abstractions.*—Stephen Jay Gould, *Full House* (2011, p. 41)

A key moment in the work of philosopher Gilles Deleuze is his confrontation with Nietzsche's claim to be "overturning" Plato's philosophy (Deleuze 1983, 1990, 1994). For most interpreters, and arguably for Nietzsche, such an

<sup>3</sup> The definition of "fitness" and "inclusive fitness" is hard to grasp here, but is related to the concept of foraging efficiency: essentially, over a given period of time foraging efficiency is determined by how many times the robots are successful at pushing food to the wall. An individual's inclusive fitness (their own success at transporting food + the food shared with them) determined the probability of their genome being transmitted to the next generation.

<sup>4</sup> There is some slippage between the terms "real" and "physical"—often the scientific publications in question use the two terms interchangeably. I follow their usage where possible, otherwise default to "physical" to refer to robots that are extended in space, use electricity and are built out of plastic, metal and other materials.

inversion concerns the opposition between the transcendent realm of the Ideas, and the fallen or immanent realm of copies. To invert such a Platonism would be to give the immanent pride of place rather than the transcendental—to make of (material) substances the basis of our ideas instead of those ideas as the source of substance and the world. Indeed, the standard reading (expressed in the quotation above by Stephen Jay Gould) is that Plato's ideas were mystical essences that existed in a plane of reality separate from that of humans and animals and other "concrete earthly" reality. The "inversion" of Platonism in which the material or the concrete is elevated over the realm of ideas begins at the latest with Aristotle's critique of Plato, and is revived in various ways throughout the ages, but meets with its greatest success in Marx (for whose materialism Aristotle is central) and Darwin (whose reversal substitutes statistical averages for essences, and crowns this-worldly variation as sovereign). Contemporary obsession with "emergence" might be said to express something of this immanentism.

However, Deleuze harbored some skepticism about this reading of Plato (I rely here primarily on Daniel Smith's reading of Deleuze). Deleuze's interpretation is refreshingly non-dualistic—it is not the case for him that Plato is distinguishing between a transcendent realm of ideas and a fallen realm of copies or images, but offers instead a tripartite relation amongst idea, true copy and false copy, or among idea, copy and simulacrum. Copies in Plato are not all equally subordinate to the ideas they mimic—indeed they do not mimic at all, but instead *participate* in the Idea, some much better and more authentically than others. The sculptor captures the image of a man when she understands the internal and external resemblance of the copy she creates to the model she observes; but the "simulacrum" or false copy bears only a semblance, it looks like the real thing, but isn't the same "on the inside" as it were. According to Deleuze, Plato observes the problem of the *rivalry* amongst copies, and the problem of deciding amongst them. The truest copy, the copy that truly resembles the Idea (truly participates in it) is opposed to the copy that is posing or dissembling: the false claimant. Thus for Deleuze the proper inversion of Plato would not be to elevate the fallen world of copies over that of the Idea, but an even more difficult task: to give the simulacrum—the false claimant—a positivity all its own. Thus, the question of whether robots simulate evolution, or actually participate in it (i.e. actually evolve), might be put to the test of Deleuze's reading, by asking how the scientists themselves conceive of and practice the exploration of evolution using robots. It seems clear that most scientists would count themselves alongside Darwin, as opposed to a Platonic metaphysics of ideas—but would this imply a commitment to robots as simulations, or as participants in that which they model?

#### 4 Robot trouble

Skepticism about the use of robots to study life comes naturally; certain aspects of the distinction between the natural and the artificial are reassuringly intransigent. And yet, as with anyone who studies animals, robot scientists love to chip away at those divisions. In the 1990s the field of Artificial Life famously staked out an

extreme position, with folks like Christopher Langton earnestly insisting that, because the pixelated creatures on their screens obeyed evolutionary laws, they were therefore alive. But slowly other researchers, like Rodney Brooks, were not so much convinced that robots or their simulations were alive—but that they could not be (or no longer be) controlled. In this respect they were not alive, but nor were they completely under the control of their creators—much like living things bred by humans. For most participants in the fields of robotics and artificial life, the processes of biology, evolution or the meaning of “life” were useful primarily in order to build better robots. John Holland and David Goldberg applied the ideas of evolutionary theory to the design of algorithms; researchers in robotics initiated projects in evolutionary, cognitive, developmental, affective and epigenetic computing; the “inspiration” of nature was oft cited amongst engineers as “biomimesis” or “bioinspired” technology (Asada et al. 2009; Bar-Cohen 2005; Berthouze and Metta 2005; Berthouze and Ziemke 2003; Garnier 2011; Goldberg and Holland 1988; Holland 2001, 2006; Krause et al. 2011; Lungarella et al. 2003; Mehmet Sarikaya 1995). Robots provide obvious power for understanding issues of locomotion, perception and communication, as well as more recently, emotion and human health. Within a couple of decades the successes of these allied fields were so substantial that it is now routine to talk about the inclusion of “genetic algorithms” or GAs in everything from a financial trading software to music production to (reflexively) the analysis of human or animal gene network data (Dasgupta and Michalewicz 2013).

Many historical and critical readings of this topic have grappled with aspects of the problem I address here, especially concerning the way life and information have been disentangled and re-entangled across history (Bensaude-Vincent 2009; Emmeche 1996; Helmreich 1998; Riskin 2007). Central to these discussions have been the historical and theoretical re-evaluations of cybernetics and its role as a seductive language for scientists across the spectrum of disciplines, but especially within biology and evolutionary theory (Hayles 1999; Kline 2015; Pickering 2010). A case like Grey Walter’s tortoises (discussed in detail in Pickering as a kind of “ontological theatre”) gives us a very clear glimpse of the prehistory of the kind of robots engaged by the scientists in the Hamilton’s rule experiment. Walter’s tortoise was, as he clearly labeled it, an “imitation of life” (specifically the brain), a model of the brain. It would be eminently cybernetic (as I think Pickering shows) to consider the turtle not a simulation of a brain but in fact a tool for participating in thought.

At roughly the same time as the computational embrace of evolution tightened, and often with very little interaction, evolutionary theory itself has seen a series of changes, debates, borrowings and reconfigurations—so much so that there is now a marked split between those who study *evolutionary theory*, and those who study living organisms (now or in the past). The link between evolutionary theory and the study of life has steadily become less obvious with generations of thinkers applying the concepts of evolution to everything from economic growth (Nelson and Winter 1982) to epistemology (Toulmin 1972) to universes (Smolin 1997). As “evolutionary theory” has grown, speciated, and recombined, it has come to be regarded less as a theory of life and its organization and more as an incredibly powerful



theory of change and diversity in any system. Daniel Dennett among others claims that evolutionary theory is “substrate neutral” (Dennett 1995). Even mainstream evolutionary theory in the hands of people like Richard Dawkins, E. O. Wilson, or Martin Nowak struggles with the abstraction of the theory from the empirical work of those who have long studied particular species and ecologies. Add to this the problematic insertion of the term “behavior” and the subject becomes even more diffuse: evolutionary psychology, sociobiology or cooperation and sociality in humans and animals have all come into their own by using tools and ideas of evolutionary theory not just to understand the origins of individual morphology or behavior, but those of interaction and “social” behavior as well. It is clear that evolutionary theory has taken on not just a life but an ecology of its own.

Classical biologists might be justifiably skeptical: driven to extremes, evolutionary theory of this kind loses touch with the empirical—as well as with other domains of theory like those of ecology, development and physiology. Not all biologists would give evolutionary theory such pride of place in a theory of life, and the range of novel problems and approaches that have emerged in recent decades clearly indicate that the gene-centric theory of life has run its course. Among other recent changes pointed to as evidence of a challenge to the dominance of the neo-Darwinian synthesis are metagenomics and the reorganization of the tree of life, epigenetics, lateral gene transfer, new understandings of metabolism, co-evolution, symbiosis, mutualism and parasitism. All these things may well challenge many of the established aspects of evolutionary theory—and they demand exploration, experimentation and the observation of living things (Dupré and Malley 2009; Gilbert et al. 2012; Gilbert and Epel 2009; Helmreich 2003; Landecker 2011; Sapp 2009; Wimsatt and Griesemer 2007).

But at the same time, computer simulation has risen in respectability: not only has there been a marked shift away from “law-governed” theories to “model-based” theories over the course of the twentieth century, but also a recognition that computer simulation might be scientifically and philosophically novel (Creager et al. 2007; Morgan and Morrison 1999; Grune-Yanoff and Weirich 2010; Humphreys 2004; Lenhard 2006a; Schweber and Wächter 2000). Debates have emerged about the difference between simulation and experiment, about the rise of “exploratory” experiments, about the role of surprise generated by simulations, and about the epistemological and ontological consequences of understanding based on the radically advancing computational power available to scientists (Franklin 2005; Humphreys 2008; Keller 2003; Lenhard 2006b; Parker 2009; Rohrlich 1990; Steinle 1997; Winsberg 2008). Robots, therefore, seem to fall into this niche of respectability as they too become more powerful, diverse and tractable as tools of exploration and experiment.

To be sure, robots are by no means restricted to the study of evolution—there are also robust uses of them to study human psychology, cognition and emotion, as well as more traditional issues of locomotion, vision, hearing, communication etc. Furthermore, robots can be used *in place* of animals/humans or they can be studied *in interaction* with animals/humans. Robots can be treated a bit like model organisms responding to an experimental set-up, or they can be used as traps, lures or decoys that provoke behavior or reaction from an animal or human. The

distinction between a robot that *simulates* something else (stands in for) and one that *participates* in something is not at all clear, whether to the biologists using robots or for those observing the scientists (Keller 2007; Suchman 2011; Wilson 2010).

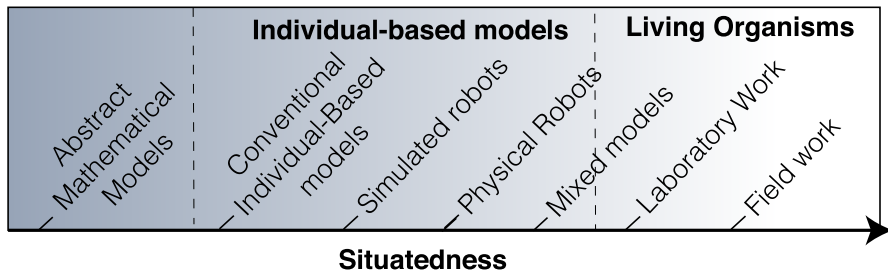
Using robots (or computer simulations) to study *evolution* is apposite though because experiments on evolution are difficult to design in the first place. The time-scales involved require experimental set-ups that accelerate time relative to that of humans and other animals, as well as a significant degree of inference and assumption. As such, experiments on evolution have typically been restricted to fast-reproducing organisms like microbes or *Drosophila* and some kinds of plants. The mathematical theory of evolution we have today thus allows considerable analytic and predictive power, but often requires an unsettling degree of simplification and only a tentative generalizability. Computer-based simulations offer a way to add “complexity” back in, and robots, therefore, seem to be the next obvious step in such an exploration.

The mathematical descriptions employed in mainstream evolutionary theory, especially those drawing on the tools of game theory, now allow an impressive level of complexity in modeling and observing simulated evolutionary dynamics within a population—but they remain computer simulations. For many scientists, the question remains: how do these simulations relate to the nature they describe? Philosophers of simulation pose this question routinely as well (Dutreuil 2014; Huneman 2014). Winsberg (among others) offers a helpful distinction by pointing out that both in simulations and in experiments, there is a difference between the *object* of explanation, and the *target* of explanation—a mouse model can be an object while the target is humans or human cancer, just as a simulation of fluid dynamics in a computer can be an object while the target is the center of a black hole—something that manifestly cannot be accessed by humans (Winsberg 2008). Target and object can be distinguished without depending on other spurious distinctions like digital/analog or real/virtual and can be distinguished equally in bench experiments in a lab, field experiments, or experiments using software models of phenomena.

Such a distinction would seem to appeal to the experimental biologists who are using robots, as it implies the possibility of a continuum of relations between object and target. Indeed, in a recent review by Mitri et al. (2013), the uses of robots in biological experiments concerning social behavior are surveyed; they lay out just such a continuum that they call the “scale of situatedness” (see Fig. 1).

On the right-hand end of the scale are versions of “situatedness” such as “field work” and “laboratory work” which “include the whole complexity of the organisms and their environment... but rarely permit the unambiguous demonstration of causations (1)” especially concerning social behavior—pure ‘target’, to be sure, but seemingly as complex and inaccessible as a black hole.

On the other end, they put “abstract mathematical models” that “boil down collective systems to their minimal components.” They are abstractions and formulas that model populations as a whole, rather than individuals in populations. Mathematical abstractions are not necessarily dependent on a computer for their existence (which is not to say, of course, that some humans might not increasingly depend on them to calculate or keep track of the abstraction)—their power comes



**Scale of situatedness: the extent to which individuals are embedded in an environment that they can sense and modify. Based on a diagram in Mitri et. al. 2013**

**Fig. 1** The scale of situateness from Mitri et al. (2013)

from representing the target with relatively simple formulas. Not all such formulas bear the relation of object and target, however. Hamilton’s rule, for instance, is less descriptive than predictive: it does not model the actual evolutionary system so much as predict an outcome based on measurements and observations.

Next on the scale are so-called individual or agent-based models that can take into account the varying behavior of a potentially large population of individuals. These models necessarily inhabit a computer simulation (rather than a formula on a page or screen in the case of ‘abstract mathematical models’), usually represented not graphically but as something like a database or spreadsheet of changing parameters over time. Such models depend on the ability to both compute the complexity of interacting individuals (each of which might have its own more or less complex genotype and phenotype) in a population and in most cases to display that computation in some form (though not always as a graphical visualization).

Further down the scale are robots (both physical and simulated), which are essentially agent-based models in physical, programmed robots that interact in real (or simulated) space. Physical robots are better than agent-based model simulations because they do not need to make simplifying assumptions about the physical environment or the physical characteristics of the agent/individual (i.e. physical organization of the robot). As the authors put it, realism in an agent-based model is *costly* and *complex* whereas “the laws of physics are included ‘for free’ in robotic models (2).”

There are several things to say about the “scale of situatedness” and the claims made for it here. Most importantly, there is both an explicit and an implicit association of “situatedness” with “embodiment”—abstract models and computer simulations are experiments in an abstract or digital space—whereas organisms are embodied and situated in real physical space. The scale is implicitly one of the complexity of bodies and their environments. Robots straddle a boundary: they share some of this bodily complexity with organisms at the same time that they share putatively simpler (or just more controllable) computational and digital modes of existence with simulated organisms. That robots have bodies is not interpreted as

a metaphor or a supplement, but instead as a crucial determinant of what robots are and can be. Embodiment is central to cognition not just in the sense that the robot needs to perceive and sense the world, but also implies that the kind of body it has will transform the kind of cognition of which it is capable.<sup>5</sup> When the authors say that the “laws of physics are included for free” they mean that all of the complexity of the robot body—its weight, orientation, speed, the distance between its eyes and its wheels, etc.—do not need to be simulated by a digital computer; they are just there, as part of what a robot is.

Deleuze’s reading of Plato bears a striking similarity to the problem of simulated and physical robots discussed here. Rather than a simple opposition between real (living things) and fake (robots), the “scale of situatedness” seems to set up a scale of rivalry along which some kinds of copies are better than others: a model organism is a true claimant to the robots’ false one, but the robot is a true claimant to the agent-based model’s false one. But as in the case of Plato’s philosophy, the question remains: on what basis does one choose one rival over others? Aristotle’s criticisms of Plato turned on this issue as well: for Aristotle, Plato’s theory of Ideas failed to adequately distinguish species and genera, leading to a fundamentally capricious system. Aristotle, rather, would insist on a clear taxonomic principle distinguishing the organism from the model organism, the robot from the organisms and robots from each other. An implicit reliance on such an Aristotelian taxonomy is precisely what troubles some biologists about using robots: *they simply are not alive*.

It would seem therefore that the question is not ‘what’s the difference between an animal and a robot?’ (model and copy), but ‘what is the difference between a physical robot and a simulated robot?’ (copy and simulacrum)? For Plato, the danger of the simulacrum is that it resembles the original on the outside, but lacks internal similarity (it is not a true copy). This would seem to describe any robot or especially classical automata that have the appearance of being what they copy but possess a radically different internal organization (the story of the mechanical Turk notwithstanding). Robots with fleshy plastic masks that simulate emotionally specific expressions are troubling and uncanny because they copy (increasingly well) the outside without being the same on the inside. Whatever we think it is that “makes us human”—it is present only in the *behavior* of the robot, not in its substance. So just what might biologists be observing when they decide to use a physical, embodied robot instead of a simulated one? That is to say, on what grounds do they distinguish these claimants to participation, and in turn, place them in order on the scale of situatedness?

<sup>5</sup> See e.g. (Chrisley and Ziemke 2006) Mitri et al. cite two influential books in cognitive science: Varela, Thompson and Rosch, *The Embodied Mind: Cognitive Science and Human Experience* and Andy Clark, *Being There: Putting Brain, Body and World together again* (Clark 1998; Varela et al. 1991). More recently, and directly relevant is the work of Josh Bongard, whose book (with Rolf Pfeifer) *How the Body Shapes the way we Think: A new view of intelligence* lays out the specifics of the “embodied turn” in both cognitive science and robotics (Pfeifer and Bongard 2007). The similarity to debates about the role of language and cognition within anthropology (Sapir and Whorf) and philosophy (Wittgenstein, Quine) is sometimes also noted, though more often this tradition is linked to phenomenology and a certain interpretation of Heidegger advanced by Hubert Dreyfus and taken up by some computer scientists and psychologists (Dreyfus 1999; Winograd 1995).

## 5 One robot may hide another

In their review, Mitri et al. present a couple of nice examples to illustrate the difference that physical embodiment makes to experimental science, and how it might serve as a basis for distinguishing along a scale of situatedness. In one study, of cockroach behavior, an agent-based model reached one outcome, while a robot model reached a different one—because in the agent-based model there was no way for one cockroach to “hide” another one—all cockroaches were visible to all other cockroaches at all times—the God’s eye view as experienced from the perspective of a cockroach. The embodied robots, however, included this feature of physics “for free” (the authors’ words) and as such the result of the physical experiment was different from that of the agent-based model inside a computer. Similarly, a study of ant-like robots’ ability to collect objects revealed that the physical ants were much less efficient than the ants in an agent-based model because they had bodies that bump into each other and slow them down. May and Schank have also demonstrated that the shape and size of a robot body are important—and in their case, their robot rats actually demonstrated something that proved to be a testable hypothesis about living rat behavior and the shape of their bodies (Schank 2004). And in other examples, the authors refer to the use of robots in “mixed-models”—that is experiments where robots are used in association with living animals in order to elicit or respond to specific behaviors—not unlike the long-standing practice in ethology of using lures or decoys, as several papers in the field like to point out.<sup>6</sup>

But what these examples do not reveal is whether there is a difference between *physical* robots and *simulated* robots. In all of the above examples, the salient difference is between an experiment using lots of robots, and one using “agents”—abstract representations of the parameters under study. But in between these two kinds of situatedness are “physics-based” robot simulations—simulations of the robots in question which include more of the physical parameters of the robot than an agent-based simulation would, but presumably less than that of the physical robot. In these physics-based simulations, the price of the laws of physics starts to fluctuate: how detailed must a physics-based simulated robot be to be the same as a physical robot? What delta is permissible between the two? At what point does it become plausible to assert that they are the same (in essence, to assert that the simulated robot “has a body”)?

There is obviously a lot of complexity to a robot simulation, and not much information is offered by the experimenters when it comes to understanding what a simulated robot is or does. The materials and methods sections in the work of Floreano and Keller, for instance, often describe in detail the robot’s configuration, precise dimensions and speed, programming, etc. as if all the robots in the study were physically real, and they rarely refer readers to the simulation tools used. Some online digging reveals that in the case of the work of Floreano and Keller, the simulation environments they use are in-house programming environments called “TEEM” (The Open Evolutionary Robotics Framework) and “Enki” (a fast 2D

<sup>6</sup> Frequently this claim relies on reference to Tinbergen (1951).

robot simulator) created by co-author Markus Waibel.<sup>7</sup> They are modular, extensible software environments of a standard sort that can be used to cobble together simulated robots and run experiments using these robots. The fact that these researchers can refer to both the physical and the simulated robots without clearly distinguishing them has at least one good reason behind it: both the simulations and the physical robots by definition use the same operating systems, programming environments and software tools and libraries—a robot might have an embedded controller with a version of Linux running on it that allows the robot to be controlled and to perform its various tasks—and that program will have been initially developed on a computer running Linux and an *emulation* of the same software and its robotic hardware environment. The relationship between the prototype and physical robot is capable of being more or less exactly replicated in the physics-based simulation because many of the software parts are interchangeable (even if it requires some admirable wizardry to make this whole system function). It is a feature of all robots that they are first simulated in software and then, essentially, “printed out” in physical form and set to work—and it is always the differences, breakdowns, or hardware-dependent surprises that form the core of an iterative process of inquiry in the robot sciences. Indeed, it shares this feature with all software programming where systems can be “bootstrapped”—the design of new software (such as an operating system) or new hardware (such as a processor) can be implemented virtually inside another by programming a simulated environment for it to exist in before it is “real” in the intended final sense. The assumption that there is no difference between a simulated and a physical robot is therefore both deeply ingrained in the practice of creating robots, and warranted by an understanding of what gives the robot its being: its software.

But to return to Deleuze, the problem of distinguishing between false claimants is that they resemble the copy on the outside but not the inside—which is to say, they appear the same in some respects, but are not motivated or actuated by the same internal workings, whatever that might be—they do not really *participate* in the idea they copy. The simulated robots, however, offer a twist on this: they may have a visual representation (such as a 3D rendering on a screen) that resembles a physical robot, but what is indistinguishable is actually the internal aspect—the program or software governing the robot. Simulated robots have simulated bodies but their “brains” and “genes” are identical to those of physical robots: they are identical on the inside, but share little or no resemblance on the outside. It is this fact, if we follow Deleuze, which serves as warrant for the scientists to treat them as interchangeable—indeed, it is not clear whether from this perspective simulated robots should appear to the left or the right of physical robots on the “scale of situatedness”—clearly, they are different, but it is not their internal resemblance which allows one to distinguish true from false claimants.<sup>8</sup>

---

<sup>7</sup> See for example, the projects available at “Teem The next generation open evolutionary framework,” available at <http://lis2.epfl.ch/resources/teem/> (Last visited Jul 23, 2017).

<sup>8</sup> To push this logic even further, it may well be possible to imagine creating, for instance, a 3D cinematic representation of a robot having a far better outward resemblance (and identical internal resemblance) of an animal or human than can be achieved with a physical, sculpted and manufactured

But if this is so, if the simulated robots can be said to be no more or less false a claimant on what they model than the real robots, then why not just run simulated robots through the motions and report on that? Implicit in the design of these experiments is that physical robots are necessary to produce that minimum of difference—that surprise—which can only come by running a “real” experiment (Rheinberger 1997). The expectation of the researchers seems to be, perhaps, that physical robots will confirm what the simulation demonstrates, but that they must be included, observed, and the hypotheses thereby tested in order to assure them that the simulation acts as the physical robots would. A good researcher would no doubt be attuned to understanding why the physical robots fail for interesting as opposed to routine reasons. However, in every case reported by Floreano and Keller so far, they use physical robots in the experiment—but only a few—and they do not seem to do anything other than confirm the results of the simulation. They produce no surprises, but are nonetheless essential to confirming that they are, as yet, unnecessary to the experimental result.

## 6 These aren’t the droids you are looking for

There are some obvious reasons why scientists might prefer robots (whether physical or simulated) to animals—they can do all kinds of ethically suspect things with robots that they cannot do with animals. The price/complexity ratio is very good and always getting better. The robots do just what they are programmed to do (except when they do not), which is useful for establishing precision in an experiment. But there are also obvious reasons why an experimenter would prefer *simulated* robots over physical ones: they can reproduce and mutate dramatically faster than the embodied robots, they do not break down as often, and they can be reprogrammed and reconfigured much more quickly.

The trade-off of course is that robots are not animals, and simulated robots are not robots. On the “scale” of situatedness, there appears to be a clear order or hierarchy. But the source of this new *scala artificio* is not entirely clear, and perhaps not entirely untheological. It relies to a large extent on our intuitions about what is real and what is not; but it is also dependent upon a certain definition of that which is authentically alive. Perhaps in the background there is an unquestioned return to a kind of vitalism; or perhaps there is a sense that it is simply “obvious” what is alive and what is not. If it were possible, given this metaphysics of situatedness, scientists would prefer to manipulate (and understand) real organisms, and barring that, real robots, and barring that simulated robots and so on. Simulated robots must exist in simulated environments and there are necessarily some kinds of physical parameters that it is either too difficult, too costly, or impossible to include—so the simulations are necessarily not identical with physical robots, even if experimenters choose to act as if they are.

---

Footnote 8 continued

robot. In this case, the simulated robot should by all rights appear to the right of the physical robot on the scale of situatedness, and not the reverse.



On the one hand, the troubling implication of such a spectrum is that mathematical abstractions are figured as the least appropriate object for understanding a target—as opposed to the traditional assumption that they bear a kind of representational accuracy, if not identity with the target (mathematics as the “language of nature”). But on the other, it seems to put mathematical abstractions, models, robots, model organisms and living organisms all on the same ontological plane—formulas are simpler versions of agent based models, which are simpler forms of robots, which are simpler forms of organisms, all of which *participate* in something not quite specified: life, evolution or social behavior. Which is to say, it collapses the object/target distinction.

Thus, there appears to be an epistemological tension here: on the one hand robots (as well as mathematical formulae!) are merely physical entities which can be simulated as other physical entities can (assuming we get the physical theories “right”) and what they do is ontologically indistinguishable from what mathematical formulae do or from what organisms do. But on the other hand is the assertion that *robots are different*: robots are *embodied* and so they will exhibit forms of behavior and/or cognition that can only come from being embodied—i.e. something that cannot be or does not reduce to their program, something that comes of having a “body” not just a “mind” (or program). This extends to the environment as well: on the scale of situatedness, fieldwork is preferred to laboratory work, and a commitment to a certain mode of realism or direct access to the real of nature is evident. Thus the choice to work with robots appears strictly pragmatic: robots give some kind of partial access to (more complex) nature in return for being more tractable (simpler). It sets up a *rivalry* between different kinds of objects which are also targets. The ordering reveals two things: an implicit order according to which the object is an authentic expression of some unstated essence: life, evolution, behavior, etc. And secondly, an order according to which the object can be controlled or manipulated in the service of modeling, understanding or confirming theories of life, evolution, behavior, etc.

The implicit definition of embodiment here, however, is a physical one. “Having a body” means “including the laws of physics for free”. By doing so we gain purchase on the physics of bodies and environments that we would otherwise have to painstakingly model in a software system. But in the case at hand what is being simulated is something else: *social behavior*, and in particular motivated social behavior that can be labeled as either selfish or altruistic. The scientists involved are no doubt committed to the epistemological claim that all such behaviors are fundamentally physical—either in a reductionist or an emergentist sense. There are genes which are biochemical entities that interact with the physical features of an environment which are then translated into proteins and on up the chain into organisms whose metabolically self-sustaining whole is liable to behave in predictable ways—such as by choosing to share food with another organism (QED, Hamilton’s rule).

The question the scientists do not pose, however, is whether these robots have a “social” body (real or simulated) as well as a “physical” one, and whether there is any difference. Mitri et al. suggest that “Robots are useful when properties of the physical environment (e.g. visual and spatial effects, friction and collision) are



likely to influence the outcome of the social behavior (8).” This would seem to imply both that social behavior is a feature of the environment and that there is a kind of social behavior that is not influenced by physical environments. Much of what these robots do, in Floreano and Keller’s experiments, is physical—concerning the physics of moving themselves with motors, perceiving ‘food’ by IR cameras and ‘seeing’ colors. But the behavior is governed by software that essentially collapses neurons and genes into one kind of thing (the genes weight the neurons which determine what the robot does or does not do). So the interaction between a robot and a physical environment or between two robots is highly deterministic, and it all comes down to the software. This in turn is another source of the warrant for treating physical and simulated robots as identical.

But is “social behavior” simply “physical behavior” governed by brains and genes? Could it be something more, and could the robots reveal it? Do the scientists expect—or perhaps hope—that the robots will surprise them? Perhaps instead of simply deciding to share or not share a food item they might stockpile it? Or decide instead to divide it and extort the other robots for it, or maybe cook it? Or conduct a ritual before consuming it? How fanciful might one’s hypotheses about sociality become when one is waiting for surprise with robots?

Unfortunately, to make things even worse, the very concept of “social behavior” (to say nothing of “altruism” or “self-interest”) is so elastic and vague that it is hard to know what most scientists mean even when humans and animals are both object and target, much less robots (Bergner 2011; Levitis et al. 2009; Piliavin and Charng 1990). One cynical answer to the question “Why use physical robots?” is simply that it is cool, or that it brings in funding, or that interdisciplinarity is (over)valued. However, the larger question concerning simulation and the “style of reasoning” that is being developed does not thereby disappear—robots and simulations have entered the practice of science to such an extent that it is no longer possible to continue to treat them as gimmicks or illustrations. They are, in increasingly many cases, the work on the basis of which scientific knowledge moves forward, develops and is argued about today—and they drive the development of technologies and infrastructures that facilitate their use for just this purpose.

But neither is it possible to treat these robots as less complex versions of living organisms. To do so is to misrecognize what they are, to mistake the false for the true claimant, and even more importantly, to follow Deleuze, to fail to give the false claimant a positivity of its own. To investigate, forthrightly, the internal difference that these robots harbor vis-à-vis animals and humans, and even, ultimately to argue against the existence of any model or original.

## 7 Do robots have a life of their own?

When we learn something from an experiment using robots, are we learning about robots or are we learning about something else (or both—one thing for the robot scientist and another for the biologist)? Floreano and Keller tell us that we are learning something about (“confirming”) Hamilton’s rule when we watch these robots do what they have been programmed to do. The algorithms of the robots

instantiated in their physical and simulated bodies unfold in accordance with our expectations about how ants or bees or humans would do the same. We are then comparing expectations to expectations and confirming their identity—this tells us that theory is correct for robots, although does not necessarily tell us that it is true for animals. It also does not necessarily tell us anything we do not already know about robots.

Given these robots and their algorithms, we can assert that simulations (including robots simulating animals) *animate* theory: they bring theories to life in time. In the case at hand the robots are animating Hamilton's rule. Eric Winsberg's term for this is that simulations are "downward"—they draw on theory and perform it in a computer rather than being the kind of thing from which one makes observations and builds theories (upwards) (Winsberg 2008). But the spatial metaphor is misleading (or relies too much on a hierarchical relationship and separation between theory and observation) because what simulations do is less about higher and lower, and more about static and dynamic. A simulation is preferable to an equation because it gives the equation life—it allows one to observe an animated version of an equation—or a whole bunch of interacting equations that transcend the capacity of humans to cogitate upon. Confirmation becomes something the simulation does for us by enacting the theories we devise. It takes a great deal of work (and an exquisite expertise in the use of such tools) for a program or simulation to fail to confirm the theories it is designed to animate. In most cases, failure to confirm is written off to a failure—or artifact—in the software or simulation system itself. Only in rare cases does a simulation produce something that might authentically be called a surprise, an insight, or an advance in the theory. In the case of the test of Hamilton's rule described here, it is safe to say the authors do not report any such insight; though the possibility is not thereby ruled out.

If equations can be observed and experienced, then they have the capacity to generate insight, surprise and difference in an experimental system. They do in fact possess the potential to be "upward" in Winsberg's sense, but only if one approaches them differently—as that which must be observed in its own right: not a true copy but a false one, a simulacrum with its own positivity, its own internal difference as a motor of change and exploration. There are many examples of such simulations generating both surprise and new difference and understanding. In microcinematography, for instance, early theories of development and duration (such as that of Bergson's) were singularly suited to the medium of cinematography and allowed scientists to see something: but were they seeing life or were they seeing a simulation of life—a theory animated before their eyes? Microcinematography animated theories of life using the film camera as a way to bring cellular growth and division back to life (from the dead discrete stills and images of photographs, histograms and illustrations); and it was by watching such animations that surprise was generated and new phenomena (e.g. apoptosis, motility) explored. Similarly, L-systems, those rudimentary attempts to simulate algae growth using formal languages (primitive programming languages) animated particular theories of cell division and growth. L-systems animated a theory of life that allowed biologists to watch not a living organism but a theory of a living organism. But when observed the surprise they produced swerved away from biology and into

mathematics and computer science, where a veritable bestiary of formal languages was observed, catalogued and compared. Today almost no biologists use formal languages for the study of life—but most computer scientists and many mathematicians are steeped in their dynamics and behaviors before they even finish an undergraduate degree (Kelty and Landecker 2004; Landecker 2005, 2009).<sup>9</sup>

Simulations therefore are obviously “downward” when they animate a theory, but they can also be upward, with a swerve, when observed. They generate surprises and insights, and they may shift attention out of one domain (the biology of algae) into another entirely (the creation of formal languages). When this happens, there is no obvious commitment to realism, but there is a commitment to curiosity—the simulation does not represent something, it literally *becomes* something else. When this happens, a simulation is neither an attenuated nor a false copy of the real, but an object and process with its own complexity, an *assemblage*.<sup>10</sup> Simulations therefore (and one must agree with Winsberg here, even if the metaphor is wrong) take on a life of their own; some of them even escape and find new niches in neighboring disciplines.

To say that the robots *participate* in the experiment, however, implies that the robot (physical or simulated) is an image of something else—an instance of some other ideal form in which it participates. A mouse, as a model organism, clearly participates in ‘life’ insofar as it is a ‘copy’ of human life (or of any other animal life figured as the model or target)—but it is a true copy because (we believe) the resemblance of *life to life* (or cell to cell, or gene to gene) is what matters, and not the fact that one is small, cute, and fuzzy and the other tall, smelly, and violent.

To view robots as participating directly in life or evolution is exactly what generates the skepticism among biologists—they are so obviously false claimants to the idea of life that it is exceedingly hard to treat any experimental result as having even the faintest relevance to understanding life. These robots have “genes” and “brains” that can only be used in scare quotes: they lack that resemblance (merely homologous though it be) that we comfortably attribute to mice and men. Robots, in this sense, are mere phantasms that delude or divert us from the true forms of life.

Robots might better be understood—in cases like the test of Hamilton’s rule explored here—to participate in something other than life: they participate in *theory*. The most commonsense way to make sense of what such experiments are using robots and simulated robots to do, is that they are part of an experimental system that can, in the best of cases, reveal something about a clearly delineated phenomenon like evolution. Whether robots and simulated robots are different then takes on a more subtle and pragmatic character—akin to asking for instance, about the difference between using glass and plastic containers in an experiment

<sup>9</sup> A similar example can be found in Lenhard, “Surprised by a Nanowire,” in which he discusses the “surprise” that comes from simulating physics at the nanoscale; Lenhard tames this surprise however, by calling it “pragmatic understanding” and reducing the simulation to a kind of tool by which theories become experiments, and subsequent experiments become confirmations or falsifications (Lenhard 2006a, b).

<sup>10</sup> Smith argues that Deleuze dropped the concept of simulacrum and replaced it with that of assemblage, (Smith 2006).

(Landecker 2013) or the difference between one satellite's weather data and that of many others (Edwards 2010). Robots could therefore be understood better as part of a complex experimental system whose purpose is not to test a theory of life, but to perform that theory. As Edwards says of climate models "the main goal of analysis is not explain weather but to reproduce it." (Edwards 2010: p. 280). Robots are tools; they participate in our models of life, and our models simulate and reproduce our best understanding of life and evolution. Robots raise tricky questions about the role of physicality and embodiment in our theories of life and evolution, but they may well help solve them.

## 8 Pseudo-coda

Is that it then, have we overturned Platonism in science? We act Platonically when we presume that our models and theories, embodied and animated in physical and simulated robots, participate in what are essentially modern ideal forms: life, evolution, behavior, sociality. But it is from Deleuze we get the idea that to overturn Plato, to truly overturn this strange mysticism, would be to deny that robots, or theories, participate in these ideal forms and to instead focus on them as assemblages with their own positivity. This is not so alien an idea really, but it conflicts with some bedrock philosophy of science assumptions about the relation of theory and what it describes, or models and what they model; that discovery and justification are separable; and that ultimately the *evolution* of scientific knowledge is not a natural phenomenon independent of humans, but in a strict Kantian sense, a problem of squaring the circle of mind and world.

But the fact is that robots are out of our control in a way that is not identical to the way that living, changing, reproducing things are out of our control. Robots are simulacra with a power all their own: assemblages of software, hardware, theories of computation and embodiment, theories of evolution and neural processing, vision sensors and motors, human operators and software environments for testing, all cobbled together for a range of reasons like understanding, human companionship, capitalist efficiency, as well as insane longings and fever-fantasies of other kinds of creation, of control and of being controlled. Robots we create are not *copies* of a "wild type" of robots against which we might compare our laboratory results; there are no field studies of robots, no robot pesticides or robot tracking collars; there are no sweeping IMAX films of swarming robots on the plains of Africa. Not yet at least.

But they are part of the world now, they have been *unleashed*. This, at least, does not yet mean they are in control, but they are also quite obviously not strictly under our control. The increasing awareness that our algorithm-rich, data-intensive infrastructures interact daily—not only with humans, but also with all sorts of living creatures inside laboratories and out—means that we are challenged to develop an understanding of robots—not just of evolution, behavior, sociality—in order to get beyond Plato. Against Plato (or against "evolutionary theory" as such), perhaps we might turn instead to Hannah Arendt, and hear her call to "think what we are doing" (Arendt 1958, p. 5) as a call to think beyond robots as useful tools for animating our

theories about ourselves, and instead see them as beings that extend, or even upend, those theories.

**Acknowledgements** This paper is a revised and expanded version of a chapter with the same title that appeared in the volume *Research Objects in their Technological Setting*, edited by Bernadette Bensaude-Vincent, Sacha Loeve, Alfred Nordmann, and Astrid Schwartz (New York: Routledge, 2017). I thank the editors of this volume for their permission to publish this revised version, and for their editorial and intellectual assistance. I also thank Janina Wellman and the MECS group at Leuphana University for the invitation to present this work and to include it in this series, and the editor of HPLS and three anonymous reviewers for detailed and helpful criticisms.

## References

- Arendt, H. (1958). *The human condition*. Chicago: University of Chicago Press. [http://ucla.worldcat.org/title/human-condition/oclc/259560&referer=brief\\_results](http://ucla.worldcat.org/title/human-condition/oclc/259560&referer=brief_results). Accessed 11 January 2013.
- Asada, M., Hosoda, K., Kuniyoshi, Y., Ishiguro, H., Inui, T., Yoshikawa, Y., et al. (2009). Cognitive developmental robotics: A survey. *IEEE Transactions on Autonomous Mental Development*. <https://doi.org/10.1109/TAMD.2009.2021702>.
- Bar-Cohen, Y. (2005). *Biomimetics: Biologically inspired technologies* (Google eBook) (Vol. 2005). Boca Raton: CRC Press. <http://books.google.com/books?id=uIoDZzy-EioC&pgis=1>. Accessed 23 April 2013.
- Bensaude-Vincent, B. (2009). Self-assembly, self-organization: Nanotechnology and vitalism. *NanoEthics*, 3(1), 31–42. <https://doi.org/10.1007/s11569-009-0056-0>.
- Bergner, R. M. (2011). What is behavior? And so what? *New Ideas in Psychology*, 29(2), 147–155. <https://doi.org/10.1016/j.newideapsych.2010.08.001>.
- Berthouze, L., & Metta, G. (2005). Epigenetic robotics: Modelling cognitive development in robotic systems. *Cognitive Systems Research*, 6(3), 189–192. <https://doi.org/10.1016/j.cogsys.2004.11.002>.
- Berthouze, L., & Ziemke, T. (2003). Epigenetic robotics—Modelling cognitive development in robotic systems. *Connection Science*, 15(4), 147–150. <https://doi.org/10.1080/09540090310001658063>.
- Chrisley, R., & Ziemke, T. (2006). Embodiment. In *Encyclopedia of cognitive science*. New York: Wiley. <https://doi.org/10.1002/0470018860.s00172>.
- Clark, A. (1998). *Being there: Putting brain, body, and world together again*. Cambridge: MIT Press. <https://books.google.com/books?id=i03NKy0ml1gC>.
- Creager, A. N. H., Lunbeck, E., & Wise, M. N. (2007). *Science without laws: Model systems, cases, exemplary narratives*. Durham: Duke University Press.
- Dasgupta, D. (Dipankar), & Michalewicz, Z. (2013). *Evolutionary algorithms in engineering applications*. In D. (Dipankar) Dasgupta & Z. Michalewicz (Eds.). Heidelberg. [https://books.google.com/books?id=g4urCAAQBAJ&dq=%2522use+of+genetic+algorithms%2522+engineering+review&lr=&source=gbs\\_navlinks\\_s](https://books.google.com/books?id=g4urCAAQBAJ&dq=%2522use+of+genetic+algorithms%2522+engineering+review&lr=&source=gbs_navlinks_s). Accessed 8 June 2017.
- Deleuze, G. (1983). *Nietzsche and philosophy*. New York: Columbia University Press. [http://melvyl.worldcat.org/title/nietzsche-and-philosophy/oclc/8763853&referer=brief\\_results](http://melvyl.worldcat.org/title/nietzsche-and-philosophy/oclc/8763853&referer=brief_results). Accessed 15 October 2013.
- Deleuze, G. (1990). *The logic of sense*. New York: Columbia University Press. [http://melvyl.worldcat.org/title/logic-of-sense/oclc/19723001&referer=brief\\_results](http://melvyl.worldcat.org/title/logic-of-sense/oclc/19723001&referer=brief_results). Accessed 15 October 2013.
- Deleuze, G. (1994). *Difference and repetition*. New York: Columbia University Press. [http://melvyl.worldcat.org/title/difference-and-repetition/oclc/29315323&referer=brief\\_results](http://melvyl.worldcat.org/title/difference-and-repetition/oclc/29315323&referer=brief_results). Accessed 15 October 2013.
- Dennett, D. (1995). *Darwin's dangerous idea: Evolution and the meanings of life*. New York: Simon & Schuster. [http://ucla.worldcat.org/title/darwins-dangerous-idea-evolution-and-the-meanings-of-life/oclc/31867409&referer=brief\\_results](http://ucla.worldcat.org/title/darwins-dangerous-idea-evolution-and-the-meanings-of-life/oclc/31867409&referer=brief_results). Accessed 23 April 2013.
- Dreyfus, H. (1999). *What computers still can't do: A critique of artificial reason* (6. print.). Cambridge: MIT Press. <http://www.worldcat.org/title/what-computers-still-cant-do-a-critique-of-artificial-reason/oclc/249027531>. Accessed 29 April 2013.
- Dupré, J., & Malley, M. A. O. (2009). Varieties of living things: Life at the intersection of lineage and metabolism. *Philosophy of Theoretical Biology*, 1(e003), 1–25.

- Dutreuil, S. (2014). What good are abstract and what-if models? Lessons from the Gaia hypothesis. *History and Philosophy of the Life Sciences*, 36(1), 16–41. <https://doi.org/10.1007/s40656-014-0003-4>.
- Edwards, P. (2010). *A vast machine: Computer models, climate data, and the politics of global warming*. Cambridge: MIT Press. [http://ucla.worldcat.org/title/vast-machine-computer-models-climate-data-and-the-politics-of-global-warming/oclc/430736496&referer=brief\\_results](http://ucla.worldcat.org/title/vast-machine-computer-models-climate-data-and-the-politics-of-global-warming/oclc/430736496&referer=brief_results). Accessed 5 January 2013.
- Eiben, A. E., & Smith, J. (2015). From evolutionary computation to the evolution of things. *Nature*, 521(7553), 476–482. <https://doi.org/10.1038/nature14544>.
- Emmeche, C. (1996). *The garden in the machine: The emerging science of artificial life*. Princeton: Princeton University Press. <http://books.google.com/books?id=MU-5GAFQsIC&pgis=1>. Accessed 12 January 2011.
- Floreano, D., & Keller, L. (2010). Evolution of adaptive behaviour in robots by means of Darwinian selection. *PLoS Biology*, 8(1), e1000292. <https://doi.org/10.1371/journal.pbio.1000292>.
- Floreano, D., Mitri, S., Magnenat, S., & Keller, L. (2007). Evolutionary conditions for the emergence of communication in robots. *Current Biology: CB*, 17(6), 514–519. <https://doi.org/10.1016/j.cub.2007.01.058>.
- Franklin, L. R. (2005). Exploratory experiments. *Philosophy of Science*, 72(5), 888–899.
- Garnier, S. (2011). From ants to robots and back: How robotics can contribute to the study of collective animal behavior. In Y. Jin & Y. Meng (Eds.), *Bio-inspired self-organizing robotic systems* (pp. 105–120). Berlin: Springer.
- Gilbert, S. F., & Epel, D. (2009). *Ecological developmental biology: Integrating epigenetics, medicine*. Sunderland: Sinauer Associates. <http://books.google.com/books?id=IA5SPQAACAAJ&pgis=1>. Accessed 18 March 2011.
- Gilbert, S. F., Sapp, J., & Tauber, A. I. (2012). A symbiotic view of life: We have never been individuals. *The Quarterly Review of Biology*, 87(4), 325–341. <https://doi.org/10.1086/668166>.
- Goldberg, D., & Holland, J. (1988). Genetic algorithms and machine learning. *Machine Learning*, 3(2–3), 95–99. <https://doi.org/10.1023/A:1022602019183>.
- Gould, Stephen J. (2011). *Full house : The spread of excellence from Plato to Darwin*. Cambridge, MA: Belknap Press of Harvard University Press.
- Grune-Yanoff, T., & Weirich, P. (2010). The philosophy and epistemology of simulation: A review. *Simulation & Gaming*, 41(1), 20–50. <https://doi.org/10.1177/1046878109353470>.
- Hamilton, W. D. (1964). The genetical evolution of social behaviour. I. *Journal of Theoretical Biology*, 7(1), 1–16. [https://doi.org/10.1016/0022-5193\(64\)90038-4](https://doi.org/10.1016/0022-5193(64)90038-4).
- Hayles, K. (1999). *How we became posthuman: Virtual bodies in cybernetics, literature, and informatics*. Chicago: University of Chicago Press. [https://books.google.com/books?id=JqB6Qy9z3TcC&dq=how+we+became+posthuman&source=gbs\\_navlinks\\_s](https://books.google.com/books?id=JqB6Qy9z3TcC&dq=how+we+became+posthuman&source=gbs_navlinks_s). Accessed 8 June 2017.
- Helmreich, S. (1998). Recombination, rationality, reductionism and romantic reactions: Culture, computers, and the genetic algorithm. *Social Studies of Science*, 28(1), 39–71. <https://doi.org/10.1177/030631298028001002>.
- Helmreich, S. (2003). Trees and seas of information: Alien kinship and the biopolitics of gene transfer in marine biology and biotechnology. *American Ethnologist*, 30(3), 340–358. <https://doi.org/10.1525/ae.2003.30.3.340>.
- Holland, J. H. (2006). *Genetic algorithms and the optimal allocation of trials*. <http://epubs.siam.org/doi/abs/10.1137/0202009>. Accessed 23 April 2013.
- Holland, O. (2001). *Artificial ethology* (1. publ.). Oxford: Oxford Univ. Press. <http://www.worldcat.org/title/artificial-ethology/oclc/231909303>. Accessed 23 April 2013.
- Humphreys, P. (2004). Extending ourselves computational science, empiricism, and scientific method. New York: Oxford University Press. [http://ucla.worldcat.org/title/extending-ourselves-computational-science-empiricism-and-scientific-method/oclc/66463129&referer=brief\\_results](http://ucla.worldcat.org/title/extending-ourselves-computational-science-empiricism-and-scientific-method/oclc/66463129&referer=brief_results). Accessed 6 May 2013.
- Humphreys, P. (2008). Computational and conceptual emergence. *Philosophy of Science*, 75(5), 584–594. <https://doi.org/10.1086/596776>.
- Huneman, P. (2014). Mapping an expanding territory: Computer simulations in evolutionary biology. *History and Philosophy of the Life Sciences*, 36(1), 60–89. <https://doi.org/10.1007/s40656-014-0005-2>.



- Ieropoulos, I. (2003). Imitating metabolism: Energy autonomy in biologically inspired robots. In *Proceedings of the AISB'03, Second international symposium on imitation in animals and artifacts* (pp. 191–194). Aberystwyth, Wales.
- Ieropoulos, I., Greenman, J., Melhuish, C., & Horsfield, I. (2010). EcoBot-III: A robot with guts. In *Proceedings of the Alife XII Conference* (pp. 733–740). Odense, Denmark.
- Keller, E. F. (2003). Models, simulation, and “computer experiments.” In H. Radder (Ed.), *The philosophy of scientific experimentation*. Pittsburgh: University of Pittsburgh Press. [http://ucla.worldcat.org/title/philosophy-of-scientific-experimentation/oclc/50606388&referer=brief\\_results](http://ucla.worldcat.org/title/philosophy-of-scientific-experimentation/oclc/50606388&referer=brief_results). Accessed 6 May 2013.
- Keller, E. F. (2007). Booting Up Baby. In J. Riskin (Ed.), *Genesis redux: Essays in the history and philosophy of artificial life* (pp. 334–346). Chicago, IL: University of Chicago Press.
- Kelty, C., & Landecker, H. (2004). A theory of animation: Cells, L-systems, and film. *Grey Room*, 17, 30–63. <https://doi.org/10.1162/1526381042464536>.
- Kline, R. R. (2015). *The cybernetics moment, or, why we call our age the information age*. Baltimore: Johns Hopkins University Press.
- Krause, J., Winfield, A. F. T., & Deneubourg, J.-L. (2011). Interactive robots in experimental biology. *Trends in Ecology & Evolution*, 26(7), 369–375. <https://doi.org/10.1016/j.tree.2011.03.015>.
- Kurland, J. A. (1980). Kin selection theory: A review and selective bibliography. *Ethology and Sociobiology*, 1(4), 255–274. [https://doi.org/10.1016/0162-3095\(80\)90012-6](https://doi.org/10.1016/0162-3095(80)90012-6).
- Landecker, H. (2005). Cellular features: Microcinematography and film theory. *Critical Inquiry*, 31(4), 903–937. <https://doi.org/10.1086/444519>.
- Landecker, H. (2009). Seeing things: From microcinematography to live cell imaging. *Nature Methods*, 6(10), 707–709. <https://doi.org/10.1038/nmeth1009-707>.
- Landecker, H. (2011). Food as exposure: Nutritional epigenetics and the new metabolism. *BioSocieties*, 6(2), 167–194. <https://doi.org/10.1057/biosoc.2011.1>.
- Landecker, H. (2013). When the control becomes the experiment. *Limn*, 1(3), 6–8.
- Leigh, E. G. (2010). The group selection controversy. *Journal of Evolutionary Biology*, 23(1), 6–19. <https://doi.org/10.1111/j.1420-9101.2009.01876.x>.
- Lenhard, J. (2006a). Surprised by a nanowire: Simulation, control, and understanding. *Philosophy of Science*, 73(5), 605–616. <https://doi.org/10.1086/518330>.
- Lenhard, J. (2006a). *Simulation pragmatic construction of reality*. Dordrecht: Springer. [http://ucla.worldcat.org/title/simulation-pragmatic-construction-of-reality/oclc/209946184&referer=brief\\_results](http://ucla.worldcat.org/title/simulation-pragmatic-construction-of-reality/oclc/209946184&referer=brief_results). Accessed 6 May 2013.
- Levitis, D. A., Lidicker, W. Z., & Freund, G. (2009). Behavioural biologists don’t agree on what constitutes behaviour. *Animal Behaviour*, 78(1), 103–110. <https://doi.org/10.1016/j.anbehav.2009.03.018>.
- Lowe, R., Montebelli, A., Ieropoulos, I., Greenman, J., Melhuish, C., & Ziemke, T. (2010). Grounding motivation in energy autonomy: A study of artificial metabolism constrained robot dynamics. In *Proceedings of the Alife XII Conference* (pp. 725–732). Odense, Denmark.
- Lungarella, M., Metta, G., Pfeifer, R., & Sandini, G. (2003). Developmental robotics: A survey. *Connection Science*, 15(4), 151–190. <https://doi.org/10.1080/09540090310001655110>.
- Mehmet Sarikaya, I. A. A. (1995). *Biomimetics: Design and processing of materials*. American Institute of Physics. <http://books.google.com/books?id=BFcbAQAIAAJ&pgis=1>. Accessed 23 April 2013.
- Mitri, S., Floreano, D., & Keller, L. (2009). The evolution of information suppression in communicating robots with conflicting interests. *Proceedings of the National Academy of Sciences of the United States of America*, 106(37), 15786–15790. <https://doi.org/10.1073/pnas.0903152106>.
- Mitri, S., Floreano, D., & Keller, L. (2011). Relatedness influences signal reliability in evolving robots. *Proceedings. Biological sciences/The Royal Society*, 278(1704), 378–383. <https://doi.org/10.1098/rspb.2010.1407>.
- Mitri, S., Wischmann, S., Floreano, D., & Keller, L. (2013). Using robots to understand social behaviour. *Biological Reviews of the Cambridge Philosophical Society*, 88(1), 31–39. <https://doi.org/10.1111/j.1469-185X.2012.00236.x>.
- Morgan, M. S., & Morrison, M. (1999). *Models as mediators: Perspectives on natural and social science*. Cambridge: Cambridge University Press. <http://www.amazon.com/dp/0521655714>.
- Nelson, R. R., & Winter, S. G. (1982). *An evolutionary theory of economic change*. Cambridge: Harvard University Press. [http://books.google.com/books?id=6Kx7s\\_HXxrC&pgis=1](http://books.google.com/books?id=6Kx7s_HXxrC&pgis=1). Accessed 28 April 2011.

- Nowak, M. A., Tarnita, C. E., & Wilson, E. O. (2010). The evolution of eusociality. *Nature*, 466(7310), 1057–1062. <https://doi.org/10.1038/nature09205>.
- Okasha, S. (2001). Why won't the group selection controversy go away? *The British Journal for the Philosophy of Science*, 52(1), 25–50. <https://doi.org/10.1093/bjps/52.1.25>.
- Parker, W. (2009). Does matter really matter? Computer simulations, experiments, and materiality. *Synthese*, 169(3), 483–496. <https://doi.org/10.1007/s11229-008-9434-3>.
- Pfeifer, R., & Bongard, J. (2007). *How the body shapes the way we think: A new view of intelligence*. Cambridge MA: MIT Press.
- Pickering, A. (2010). *The cybernetic brain: Sketches of another future*. Chicago: University of Chicago Press. [https://books.google.com/books?id=1pxUngEACAAJ&dq=pickering+cybernetics&hl=en&sa=X&ved=0ahUKEwiDpYne\\_q7UAhVU5GMKHd3XAAQQ6AEILTAB](https://books.google.com/books?id=1pxUngEACAAJ&dq=pickering+cybernetics&hl=en&sa=X&ved=0ahUKEwiDpYne_q7UAhVU5GMKHd3XAAQQ6AEILTAB). Accessed 8 June 2017.
- Piliavin, J. A., & Charng, H.-W. (1990). Altruism: A review of recent theory and research. *Annual Review of Sociology*, 16, 27–65. <https://doi.org/10.2307/2083262>.
- Rheinberger, H.-J. (1997). *Toward a history of epistemic things: Synthesizing proteins in the test tube* (1st ed.). Palo Alto: Stanford University Press. <http://www.amazon.com/dp/0804727864>.
- Riskin, J. (2007). *Genesis redux: Essays in the history and philosophy of artificial life*. Chicago: University of Chicago Press.
- Rohrlich, F. (1990). Computer simulation in the physical sciences. In *PSA: Proceedings of the biennial meeting of the philosophy of science association, 1990* (pp. 507–518). <https://doi.org/10.2307/193094>.
- Sapp, J. (2009). *The new foundations of evolution: On the tree of life*. Oxford: Oxford University Press. <http://books.google.com/books?id=JuNK5bTXexcC&pgis=1>. Accessed 12 January 2011.
- Schank, J. C. (2004). A biorobotic investigation of Norway rat pups (*Rattus norvegicus*) in an arena. *Adaptive Behavior*, 12(3–4), 161–173. <https://doi.org/10.1177/105971230401200303>.
- Schweber, S., & Wächter, M. (2000). Complex systems, modelling and simulation. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 31(4), 583–609. [https://doi.org/10.1016/S1355-2198\(00\)00030-7](https://doi.org/10.1016/S1355-2198(00)00030-7).
- Smith, D. W. (2006). The concept of the simulacrum: Deleuze and the overturning of Platonism. *Continental Philosophy Review*, 38(1–2), 89–123. <https://doi.org/10.1007/s11007-006-3305-8>.
- Smolin, L. (1997). *The life of the cosmos*. New York: Oxford University Press. [http://ucla.worldcat.org/title/life-of-the-cosmos/oclc/35033598&referer=brief\\_results](http://ucla.worldcat.org/title/life-of-the-cosmos/oclc/35033598&referer=brief_results). Accessed 23 April 2013.
- Steinle, F. (1997). Entering new fields: Exploratory uses of experimentation. *Philosophy of Science*, 64, S65–S74. <https://doi.org/10.2307/188390>.
- Suchman, L. (2011). Subject objects. *Feminist Theory*, 12(2), 119–145. <https://doi.org/10.1177/1464700111404205>.
- Tinbergen, N. (1951). *The study of instinct*. Oxford: Clarendon Press. [http://ucla.worldcat.org/title/study-of-instinct/oclc/766056&referer=brief\\_results](http://ucla.worldcat.org/title/study-of-instinct/oclc/766056&referer=brief_results). Accessed 29 April 2013.
- Toulmin, S. (1972). *Human understanding: Vol. 1-*. Oxford. <http://www.worldcat.org/title/human-understanding-vol-1/oclc/469358599>. Accessed 23 April 2013.
- Varela, F. J., Thompson, E., & Rosch, E. (1991). The embodied mind. *Revue Internationale de Philosophie*. <https://doi.org/10.1111/j.1468-0149.1965.tb01386.x>.
- Waibel, M., Floreano, D., & Keller, L. (2011). A quantitative test of Hamilton's rule for the evolution of altruism. *PLoS Biology*, 9(5), e1000615. <https://doi.org/10.1371/journal.pbio.1000615>.
- Wilson, E. (2010). *Affect and artificial intelligence*. Seattle: University of Washington Press. [http://ucla.worldcat.org/title/affect-and-artificial-intelligence/oclc/557402981&referer=brief\\_results](http://ucla.worldcat.org/title/affect-and-artificial-intelligence/oclc/557402981&referer=brief_results). Accessed 6 May 2013.
- Wimsatt, W. C., & Griesemer, J. R. (2007). Reproducing entrenchments to scaffold culture: The central role of development in cultural evolution. In R. Sansom & R. N. Brandon (Eds.), *Integrating evolution and development: From theory to practice* (pp. 227–323). Cambridge, MA: MIT Press.
- Winograd, T. (1995). Heidegger and the design of computer systems. In A. Feenberg (Ed.), *Technology and the politics of knowledge* (pp. 108–127). Bloomington: Indiana Univ. Press. <http://www.worldcat.org/title/technology-and-the-politics-of-knowledge/oclc/260169089>. Accessed 29 April 2013.
- Winsberg, E. (2008). A tale of two methods. *Synthese*, 169(3), 575–592. <https://doi.org/10.1007/s11229-008-9437-0>.